DOCUMENT RESUME

ED 070 661 24 SE 015 470

AUTHOR	Voelker, Alan M.; Harris, Margaret L.
TITLE	Measuring Science Concept Attainment of Elementary School Boys and Girls.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO	WRDCCL-TR-197
BUREAU NO	BR-5-0216
PUB DATE	Nov 71
CONTRACT	OEC-5-10-154
NOTE	35p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Classification; Conceptual Schemes; *Educational Research; *Elementary School Science; *Evaluation; Science Education; Scientific Concepts; *Test Construction; Tests
IDENTIFIERS	Research Reports

ABSTRACT

Test items were developed for assessing the mastery of 30 selected science concepts on classification. These concepts were drawn from the areas of physical, biological, and earth sciences. A schema of twelve test items was developed for each concept. Procedures used in the construction and revision of these test items are described. The tests were given to beginning sixth grade children and the publication includes most of the statistical data. Separate data analyses are presented for boys and girls. (PS)

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED 070661

TECHNICAL REPORT NO. 197

MEASURING SCIENCE CONCEPT ATTAINMENT OF ELEMENTARY SCHOOL BOYS AND GIRLS

REPORT FROM THE PROJECT ON A STRUCTURE OF CONCEPT ATTAINMENT ABILITIES

FILMED FROM BEST AVAILABLE COPY

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR COGNITIVE LEARNING

THE UNIVERSITY OF WISCONSIN
Madison, Wisconsin

U. S. Office of Education
Center No. C-03
Contract OE 5-10-154

SE 015 470

RESEARCH
COGNITIVE LEARNING
DEVELOPMENT

ED 070661

Technical Report No. 197

MEASURING SCIENCE CONCEPT ATTAINMENT OF ELEMENTARY SCHOOL BOYS AND GIRLS

By Alan M. Voelker and Margaret L. Harris

Report from the Project on
A Structure of Concept Attainment Abilities
Robert E. Davidson, Lester S. Golub, Herbert J. Klausmeier,
Thomas A. Romberg, B. Robert Tabachnick, Alan M. Voelker
Principal Investigators

and
The Quality Verification Program
Mary R. Quilling, Director

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

November 1971

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on a knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Project on the Structure of Concept Attainment Abilities in Program 1 and from the Quality Verification Program. The Quality Verification Program assisted in developing tests to measure concept achievement and identifying reference tests for cognitive abilities, while the Concept Attainment staff took primary initiative in identifying basic concepts in science at intermediate grade levels. The tests will be used to study the relationships among cognitive abilities and learned concepts in various subject matter areas. The outcome of the Project will be a formulation of a model of structure of abilities in concept attainment in a number of subjects, including mathematics, social studies, and language arts, as well as science.

Contents

		Page
	List of Tables	vii
	Abstract	ix
I	Introduction	1
II	Procedures	3
	Concept Selection	3
	Test Development	3
	Readability	4
	Validity	5
	Content Validity	5
	Reliability	5
	Item Revision	5
	Subjects	7
	Test Administration	8
	Data Analysis	9
III	Results and Discussion	11
	Means	11
	Tests of Concept Attainment	11
	Tests of Task Attainment	12
	Standard Deviations	13
	Reliabilities	13
	Item Indices	13
	Item Difficulty	14
IV	Conclusions	27
	References	29

List of Tables

Table		Page
1	Lists of Science Concepts by Topical Area	4
2	Mean Scores and Standard Deviations on Lorge-Thorndike Intelligence Test and Iowa Tests of Basic Skills for Students in Population and Samples	7
3	Distribution of Fathers' Occupations for Students in the Samples	8
4	Means, Standard Deviations, and Hoyt Reliabilities for Tests of Concept Attainment—Science	12
5	Means, Standard Deviations, and Hoyt Reliabilities for Tests of Task Attainment—Science	13
6	Item Indices Based on Concept and Task Criterion Scores—Science	15

Abstract

The procedures employed in constructing and revising 12-item tests of concept attainment for 30 selected science concepts are described. The total score and individual item statistics for data collected on a group of beginning sixth grade children are presented and discussed. Separate data analyses are presented for the boys and the girls.

I
Introduction

The project entitled "A Structure of Concept Attainment Abilities" (to be referred to as the CAA Project) has as its primary objective the formulation of one or more models or structures of concept attainment ability. The development of the model or models is to be based on the collection of actual data.

In the fields of both educational psychology and science education, much discussion in the literature has been devoted to classifying and/or defining types of concepts. According to Klausmeier, Harris, Davis, Schwenn, and Frayer (1968), concepts may be defined in one or more of four ways:

1. Structurally, in terms of perceptible or readily specifiable properties or attributes,

2. Systematically, in terms of synonyms or antonyms,

3. Operationally, in terms of the procedures used to distinguish the concept from other concepts, or

4. Axiomatically, in terms of logical or numerical relationships.

Bourne (1966, p. 1) states that "a concept exists whenever two or more distinguishable objects or events have been grouped or classified together and set apart from objects on the basis of some common feature or property of each." A comparison of this definition with the definitions stated above reveals a marked similarity between it and a concept defined structurally.

The roots of both these definitions seem to be based in the scientific process of classification. As such, concepts which are formulated as a result of this classification process might be identified as members of a group of concepts known as classificatory concepts. It is concepts of this variety with which this project is concerned. A classificatory concept is defined as a concept which has three characteristics:

1. There is more than one example of the concept,

2. The properties (attributes) of the concept can be described, and

3. The concept can be labeled (named) by a word or compound word.

This definition served as the basis for the selection and analysis of the subject-matter concepts.

The major steps in pursuing the objective for formulating models for the attainment of classificatory concepts are as follows:

1. To identify basic concepts in language arts, mathematics, science, and social studies appropriate at the fourth-grade level,

2. To develop tests to measure achievement of these concepts,

3. To identify reference tests for measuring cognitive abilities, and

4. To study the relationship among learned concepts in these four subject matter fields and the identified cognitive abilities.

In the assessment of concept attainment, the learner may be asked to demonstrate his attainment of a concept by performing a wide variety of tasks. The ability to perform one or more tasks is usually taken as the criterial evi-

dence that a student does or does not understand a concept. Because of the nature of this project, it was necessary to have agreement on the nature of the tasks that would be employed for making assessments about the level of attainment of the selected concepts. In this way this variable would be controlled among the subject matter disciplines represented in terms of the construction of instruments for measuring concept attainment. The schema selected to fulfill this requirement was developed by Frayer, Fredrick, and Klausmeier (1969). The "Schema for Testing the Level of Concept Mastery" consists of 13 types of questions, each of which purportedly gives an indication of some level of understanding of a concept. Each task elicits a different type of performance from the examinee. There were 12 of these 13 kinds of questions selected for use in the study.

1. Given the name of an attribute, select an example of the attribute.

2. Given an example of an attribute, select the name of the attribute.

3. Given the name of a concept, select an example of the concept.

4. Given the name of a concept, select a nonexample of the concept.

5. Given an example of a concept, select the name of the concept.

6. Given the name of a concept, select the relevant attribute.

7. Given the name of a concept, select the irrelevant attribute.

8. Given the definition of a concept, select the name of the concept.

9. Given the name of a concept, select the definition of the concept.

10. Given the name of a concept, select the supraordinate concept.

11. Given the name of a concept, select the subordinate concept.

12. Given the names of two concepts, select the relationship between them.

From the nature of the tasks identified above, it is obvious that for the purposes of this study a multiple-choice format which requires an examinee to select an answer was preferable to a testing format which requires an examinee to produce an answer. The schema, however, is sufficiently flexible to permit either of these kinds of data gathering procedures to be utilized.

It is thus the purpose of this technical report to provide a description of the test development effort for the science concepts to be included in the study. It is a report of one aspect of step two in the overall study.

II
Procedures

Concept Selection

As indicated earlier, one of the basic prerequisites to conducting this study was to have all concepts be representative of a general class of concepts. The class of concepts chosen was identified as classificatory concepts, those which can be represented by a single or a compound word. The second prerequisite to identifying the concepts for inclusion in the study was that they should measure the level of mastery of concepts commonly taught in the elementary school at the intermediate grade levels. The third prerequisite to selecting the concepts was that they should represent major topical areas pertinent to the subject-matter discipline. Three organizational patterns for selecting topical areas were considered. The one that lent itself most appropriately for use in this study was to group the concepts into earth science, biological science, and physical science areas. A more detailed description of the rationale and procedures for selecting the concepts can be found in Working Paper No. 57 (Voelker, Sorenson, & Frayer, 1971).

The initial lists of concepts from each of the three topical areas are presented in Table 1. A total of 30 concepts was selected for inclusion in the study by randomly sampling 10 concepts from each of the three topical areas. The 30 concepts for which tests of concept attainment were developed are identified with an asterisk.

For each of these 30 concepts, a comprehensive concept analysis was completed. The framework for conducting the analysis can be found in Working Paper No. 16 (Frayer et al., 1969). The information produced by a concept analysis includes:

1. Relevant and irrelevant attributes of the concept,
2. Supraordinate, coordinate, and subordinate concepts,
3. Criterial attributes of the concept,
4. A definition of the concept,
5. Examples and nonexamples of the concept, and
6. Relationships between and among the concept and other concepts.

A detailed analysis of the 30 concepts included in this study can be found in the appendix of Working Paper No. 57 (Voelker et al., 1971).

Test Development

The information provided by the concept analyses served as the basis for constructing the test items to be used in data collection. For each of the 30 concepts included in the study, a 12-item test was constructed, each of the 12 items per test corresponding to one of the tasks of the schema. This activity resulted in the production of 360 items for use in assessing science concept attainment.

If one studies the tasks being used to measure understanding of a concept, it can readily be seen that more than one item can be generated for some of the tasks. For example, a Task 1 type item (given the name of an attribute, select an example of the attribute) could be constructed to measure understanding of each of the relevant attributes of a concept. It was decided for this project to construct just one multiple-choice item for each task for each concept. This made it necessary to have bases for making choices when such choices were necessary. These bases consisted of principles for selecting attributes, relationships, incor-

Table 1

Lists of Science Concepts by Topical Area

Biological Science	Earth Science	Physical Science
Adaptation	Air Pressure	Burning
Amphibian	Atmosphere	Condensation
Animal	*Cloud	*Conductor
*Bird	*Core	Contraction
Brain	Crust	Degree
*Cell	*Fossil	Dissolve
Eardrum	*Glacier	*Evaporation
Environment	Igneous rock	*Expansion
*Fish	Magma	Force
*Heart	Mantle	*Friction
Hibernate	Metamorphic rock	Fuel
*Invertebrate	*Meteor	Gas
*Lens - eye	Meteorite	*Liquid
Ligament	Mineral	Magnet
*Lungs	*Moon	Matter
*Mammal	Orbit	*Melting
*Muscle	*Planet	Molecular movement
Nervous system	Season	*Molecule
Optic nerve	*Sedimentary rock	Non-conductor
Plant	Solar system	*Solid
*Pore	Star	*Sound
Reptile	Sun	Temperature
Retina	*Volcano	*Thermometer
Sense	Weather	Work
Skeleton	*Wind	
Survival		
Vertebrate		
Water		

* Indicates that a test was developed and administered for this concept.

rect choices, etc. A discussion of such bases can be found in "A Structure of Concept Attainment Abilities: The Problem and Strategies for Attacking It" (Harris, Harris, Frayer, & Quilling, in press).

The initial draft of each item was prepared by a science education specialist and reviewed by the principal investigator for science. When agreement on the appropriateness of the items was reached between these two parties, the items were further reviewed by a group of graduate students in science education, each a specialist in biological science, earth science, and/or physical science. Their suggestions were considered in making further revisions in the items. These revised items were then critiqued by a committee composed of item writers from each of the four subject matter areas being studied (science, social studies, language arts, and mathematics), an experienced elementary school teacher with a specialty in reading and a measurement specialist. The final critique was conducted by the principal investigator for science and a measurement specialist. Major concerns in the item construction process were readability of the items, validity, and reliability.

Readability

Each item was specifically constructed to minimize the chance that a student would be unable to answer a question because of his inability to read the item. Care was taken to use the simplest possible language and

still be scientifically accurate at the level of a child in the intermediate grades. Some assistance in meeting this criterion of readability was obtained from the analysis of the concepts. All attributes, examples, nonexamples, and the concept definition were stated in terms that a fifth-grade child could be expected to read and understand.

Pilot studies were conducted to determine whether children of the respective age level could recognize the concept labels on the basis of their ability to read the words aloud and explain whether they knew something about the concepts. The evidence obtained from this pilot study and the independent review of sample items by outside consultants indicated that there would be little if any reading problem with the items, and that the concern for administering them in any way other than the standard one in which the students read the items themselves was not justified.

Validity

The content validity of each of the items was of immediate concern during item construction; aspects of construct validity were to be probed later using duplicate test construction, simplex analyses, and factor analyses of the results obtained using the content-valid items constructed.

Content Validity

Each item was constructed to meet the content and task specifications set for it. The schema adopted for use in measuring concept attainment specified the nature of the task to be performed by the student for a respective item. The content for each item was specified as a result of the prior concept analysis. It is recognized that the content specifications were not as precise as the task specifications because of the necessity of choosing a single attribute to be tested and the necessity of selecting incorrect alternatives for use in the construction of multiple-choice questions, for example. However, systematic construction of alternate choices was used whenever possible.

To further insure the content validity of the items, two persons who were familiar with the schema for assessing concept attainment but who were not involved in the item development process classified five random sets of 72 items, the items for six concepts in each set, according to content and task. Concept analyses were available to these persons at that time. They were able to correctly classify all but a few of the items. Any questions raised by this process were mutually resolved with the subject matter principal investigator, the measurement specialist, and the reviewer.

Reliability

Utilizing the schema as the basis for preparing the 30-concept test provided a control over the kinds of items constructed. This control over the nature of the items resulted in a 12 (tasks) by 30 (concepts) matrix consisting of a score for each of the 360 items. In addition to a single test score based on the 360 items, it is possible to obtain two other types of test scores from a completely crossed design of this nature. One alternative is to have 30 concept test scores based on 12 item tests. The other alternative is to have 12 task type scores based on items selected for each of the 30 concepts. The total score for all 360 items was rejected since it assumes that neither task nor concept variation is present. Rather than make a choice between the other two alternatives, both were done. (The theoretical problem of how to item analyze a completely crossed design of this variety remains to be solved.)

The major concern was that the reliability for the concept test scores and the task test scores be sufficiently high to warrant further study using these data. It was recognized that there may be some contradiction in what was attempted. The items were constructed to comply with the completely crossed design, 30 concepts by 12 tasks. One major objective of the project is to determine the dimensionality of the selected science concepts. If either or both of these were not unidimensional, then an internal consistency reliability estimate based upon items measuring aspects from the multidimensions would reflect this; the more dimensions present and the more uncorrelated they are, the lower the internal consistency estimate. Recognizing this and not being able to study the dimensionality of the two modes (concepts and tasks) until after the items were developed, pilot studies were conducted using the items for all the concepts for the 12 tasks. As will be pointed out later, evidence indicates that sufficiently reliable scores can be obtained for both task scores and concept scores.

Item Revision

A pilot study was conducted to gather data

to be utilized in revision of the items and the tests prior to conducting the actual study. A total of five 72-item tests was administered to every subject. Each test consisted of the 12 items for six of the 30 concepts. Items were randomly assigned to position on the tests to minimize a probable learning effect that might have taken place had the items been presented in order on individual concept tests. (If one looks at the ordering of the schema, it is obvious what is meant by the potential for the learning effect when the items for individual concepts are presented in sequence.)

The pilot study activities were conducted during January, April, and May of 1970 in Sussex and Brookfield, Wisconsin. One of the five test instruments was administered in Sussex and the remaining four test instruments were administered to children in Brookfield. The Brookfield testing was divided into two periods during which two of the four instruments were administered. Approximately 100 students took each of the five concept tests. Data collected from this pilot study were then utilized to revise the test items. However, the students in the Brookfield groups were predominantly at the 90th percentile on the Iowa Tests of Basic Skills (General), which resulted in skewed results on the tests. Therefore, compensations had to be made when utilizing the results of the tests in revising the items. In addition to examining the results on individual items, the reliabilities were computed for concept tests and task tests to be used in revising items.

The scores on the tests were subjected to analysis by the Generalized Item Analysis Program (GITAP) (Baker, 1969). The output from this program provides the Hoyt reliability and standard error of measurement for the total test scores, and in addition provides four kinds of information about the individual test items:

1. The proportion responding to each possible item response. The proportion of students who respond correctly to an item is an index of the difficulty level of that item. The greater the value of this difficulty index the easier the item.

2. The item-criterion biserial correlation. The biserial correlation coefficient is an index of the discriminating ability of the item choice. For these analyses the criterion ability used was total concept or total task score.

3. X_{50}. X_{50} is the point on the criterion scale given in standard deviation units corresponding to the median of the item characteristic curve. It is the point at which subjects with that score have a 50-50 chance of choosing that response.

4. Beta. Beta is the reciprocal of the standard deviation of the item characteristic curve at the X_{50} point. It is an index of the discrimination power of the item.

As indicated earlier, there is no known way to item analyze completely crossed data such as are produced in the design of this study. Therefore, the items were analyzed as part of a concept test score and as part of a task test score. This raises questions as to the interpretation of such results. The main referents used for interpreting the results and as a basis for making item revisions were the results obtained from the analyses of the concept scores. The tasks were fixed, and thus any arbitrary decisions were made in regard to appropriate content for incorrect choices, etc. The usual standards for item indices were not strictly adhered to, as a unique design for item analysis was being dealt with and a major object of the project is to study the dimensionality of the concepts and of the tasks. If high discrimination indices were demanded, the dimensionality may have been affected by making the items more homogeneous. Also, no attempt was made to manipulate the difficulty level of the items, since another objective of the project is to determine if any differential levels of difficulty or complexity exist in the concepts and in the tasks. Therefore, the item analysis results were used as a very general guide to help in determining whether there were "hidden" weaknesses, clues, and/or incongruities in the items and even in a more general sense to show that what we were attempting to do was possible—sufficiently reliable concept and task scores could be attained when using this completely crossed design.

The general criteria (guidelines) used in item revision were as follows:

1. Biserial r: A high positive biserial r is desired for a correct choice and a high negative biserial r is desired for the incorrect choices. The general level of biserial r for the task scores was greater than or equal to .3 for the correct choice and less than or equal to a -.3 for the incorrect choices. For the concept test scores the limit on the biserial r was increased.

Table 2

Mean Scores and Standard Deviations on Lorge-Thorndike Intelligence Test and Iowa Tests of Basic Skills for Students in Population and Samples

Test		Population	Boys	Girls
Lorge-Thorndike Intelligence	$\bar{X}$	106.6	106.11	112.23
	s		14.82	13.37
	N	2605	161	239
Iowa Basic Skills				
Vocabulary	$\bar{X}$	5.53	5.54	5.88
	s		1.41	1.33
	N	2520	181	246
Reading Comprehension	$\bar{X}$	5.44	5.29	5.97
	s		1.51	1.35
	N	2520	181	247
Language Skills	$\bar{X}$	5.24	5.04	5.82
	s		1.44	1.34
	N	2520	181	248
Work-Study Skills	$\bar{X}$	5.46	5.41	5.86
	s		1.30	1.18
	N	2520	181	248
Arithmetic Skills	$\bar{X}$	5.05	5.08	5.35
	s		.96	1.00
	N	2520	181	247
Composite	$\bar{X}$	5.35	5.27	5.77
	s		1.17	1.11
	N	2520	181	245

2. Beta: A high positive beta is desired for the correct item choice and high negative betas are desired for the incorrect item choices. When beta is equal to zero

   a. no one took the item choice,

   b. the biserial r is greater than or equal to 1.00, and beta cannot be computed but can be interpreted as approaching infinity (this is a peculiarity of the GITAP program), and

   c. beta is really zero and does not discriminate.

The revised items can be found in Working Paper No. 58, "Items for Measuring the Level of Attainment of Selected Classificatory Science Concepts by Intermediate Grade Children" (Voelker & Sorenson, 1971).

Subjects

The revised items were administered to a group of beginning sixth-grade children in Madison, Wisconsin, during the fall of 1970. The population consisted of all beginning sixth graders attending Madison public schools. A sufficient number of students was randomly selected and invited to participate in the study to result in approximately 200 boys and 200 girls taking the revised tests. The actual number of students who completed the test and for which test results were usable was 259 girls and 186 boys. Each student who participated in the study was paid a fee.

Data from the Lorge-Thorndike Intelligence Test and the Iowa Tests of Basic Skills for the

Table 3

Distribution of Fathers' Occupations for Students in the Samples

	Girls	Boys
00. Accountant	4	7
01. Architect	3	2
02. Dentist	3	1
03. Engineer	10	7
04. Lawyer, Judge	6	2
05. Clergyman	--	3
06. Doctor	12	3
07. Nurse	--	--
08. Teacher, Professor	20	15
09. Other Professional	26	15
11. Farmer	--	--
21. Owner of Business	4	2
22. Manager, Official	28	13
31. Bookkeeper	--	--
32. Receptionist	1	--
39. Other Clerical	6	4
49. Salesman	27	24
51. Craftsman, Skilled Worker	39	22
52. Foreman	--	2
53. Armed Services - Officer	--	1
54. Armed Services - Enlisted	--	1
61. Truck Driver	5	4
62. Operative in Factory	16	11
69. Other Operative	12	12
71. Fireman	2	2
72. Policeman	2	4
73. Other Protective Service	3	--
74. Practical Nurse, Nurse's Aide	1	1
75. Private Household Worker	--	--
79. Other Service Workers	14	16
81. Non-farm Laborer	3	2
82. Farm Laborer	1	--
91. Not presently in labor force	6	6
99. Not ascertained	12	10

boys and girls who participated in the study as well as for the population from which they were drawn are presented in Table 2. In addition to data on the individual subjects themselves, data regarding the nature of their fathers' occupations are presented in Table 3.

Test Administration

The testing program was carried on in two situations. Since a large percentage of the sixth grade students in the Madison public schools attend one of three middle schools, it was decided to test the selected students from these schools in their own buildings after school hours. The remaining students who did not attend one of these three schools were tested in three consecutive Saturday morning sessions at centrally-located Madison schools. The test instruments were administered in 2 1/2-hour testing sessions. During a given test session each student received a science test consisting of 72 items which took approximately 1 hour to complete and a similar test from another subject matter area. There was approximately a 1/2-hour break between taking the science test and the second test.

Data Analysis

The data obtained from the administration of the tests according to the previously described procedures produced two types of information for both the boys and the girls—separate analyses for boys and girls were conducted throughout. For each of the 30 concept tests and the 12 task tests a reliability estimate was computed, and an item analysis was conducted for each of the 360 items. Each item was analyzed as part of two test scores, the concept test score and the task test score. Total test score information and the individual item data were examined in detail according to the procedures described on page 6-7.

III
Results and Discussion

The test data from the administration of the revised items to a group of beginning sixth-grade students as previously described were subjected to analysis with the GITAP program. In keeping with the design of the study separate analyses were performed for the boys and the girls. Tables 4 and 5 present the means, standard deviations, and Hoyt reliabilities for the concept test scores and the task test scores. All concept test scores are based on 12 items, while the task test scores are based on the analysis of 30 items.

Means

Tests of Concept Attainment

The mean scores attained by the boys on the 30 tests of concept attainment ranged from 6.62 to 9.61 out of a possible attainable score of 12.00 (Table 4). For the tests in both the biological and the physical science areas, the difference between lowest and highest mean scores was approximately two points, whereas in the earth science area, there was less than a 1-point difference between the lowest and the highest mean score. Overall, the highest mean scores were obtained in the biological science area and the lowest mean scores in the physical science area.

The range of mean scores for the girls on the 30 concept attainment tests was in excess of 4 points, from 6.04 to 10.44. The pattern of mean scores for the concept tests within the specific science areas was similar to that observed for the boys. The greatest differences between the lowest and highest mean scores occurred in the biological and physical science areas. Also, the highest means were obtained for concept tests in the biological science area and the lowest means were obtained for the concept tests in the physical science area.

The girls attained a higher mean score than the boys on 25 of the 30 concept tests. This overall pattern of mean scores for girls being greater than that for boys was also noted in each of the three specific science areas; girls had higher mean scores on eight or more of the concept tests in each of the areas. (Note at this point that no consideration has been given to differences in test reliabilities for the boys and the girls. Note also that the tests have only 12 items.)

An examination of Table 4 reveals the following:

1. For both the boys and the girls the concepts for which the highest mean scores were attained were Mammal and Fish, both from the biological science area. The four concept tests on which the subjects earned the lowest scores were also identical for both the girls and the boys: Invertebrate, Cell, Molecule, and Conductor.

2. The ten concept tests with highest ranked mean scores included nine of the same concepts for the girls and the boys. Five of these concepts were from the biological science area and three from the earth science area. Eight of the ten concept tests with the lowest ranked mean scores were the same for the girls and the boys. Four were from the physical science area and three from the biological science area.

There was a difference of four to eight positions in the rank order of the mean scores on seven of the 30 concept attainment tests. The mean test scores for boys were at a higher rank order for five of these seven. There was one biological science concept in this group of ranked differences, the mean score for girls

Table 4
Means, Standard Deviations, and Hoyt Reliabilities for
Tests of Concept Attainment—Science

Concept	Mean		Standard Deviation		Hoyt Reliability	
	Boys	Girls	Boys	Girls	Boys	Girls
1. Bird	8.88	9.43	2.20	1.88	.62	.55
2. Cell	7.33	7.29	2.50	2.28	.61	.54
3. Fish	9.42	10.08	2.30	1.86	.71	.65
4. Heart (Human)	8.79	9.36	2.63	2.39	.74	.72
5. Invertebrate	7.40	7.42	2.79	2.63	.73	.69
6. Lens (Eye)	7.87	8.08	2.47	2.19	.68	.60
7. Lungs	8.95	9.45	2.82	2.61	.79	.78
8. Mammal	9.61	10.44	2.48	2.11	.76	.76
9. Muscle	8.07	7.99	2.52	2.69	.67	.70
10. Pore (Skin)	8.26	8.85	2.68	2.73	.72	.77
11. Cloud	8.22	8.67	2.58	2.03	.72	.58
12. Core (Earth)	8.68	8.99	2.66	2.22	.75	.66
13. Fossil	8.81	9.36	2.46	2.08	.70	.62
14. Glacier	8.32	8.69	2.48	2.41	.66	.67
15. Meteor	7.65	7.67	2.76	2.43	.72	.64
16. Moon	8.57	8.34	2.76	2.85	.76	.78
17. Planet	8.32	8.67	2.43	2.36	.68	.68
18. Sedimentary Rock	7.91	8.75	2.68	2.50	.71	.72
19. Volcano	8.78	9.27	2.33	2.08	.65	.60
20. Wind	8.76	9.56	2.56	2.19	.71	.67
21. Conductor	6.62	6.04	2.73	2.60	.68	.66
22. Evaporation	7.99	8.29	2.71	2.53	.71	.67
23. Expansion	7.51	7.80	2.74	2.77	.71	.73
24. Friction	7.69	7.49	2.35	2.16	.62	.52
25. Liquid	8.89	9.22	2.35	2.29	.67	.68
26. Melting	7.75	8.37	2.40	2.16	.65	.62
27. Molecule	6.62	6.99	2.48	2.33	.60	.56
28. Solid	8.58	9.56	2.76	2.21	.77	.69
29. Sound	8.16	8.54	2.49	2.27	.69	.66
30. Thermometer	8.34	8.68	2.54	2.07	.71	.57

receiving the higher rank. In both the physical science and the earth science areas, there were three concepts with wide differences in the rank orders of the mean scores. Mean test scores for the boys received the higher rank order for all three physical science concepts and two of the three earth science concepts.

Tests of Task Attainment

Data were also analyzed in terms of the mean scores earned for each of the 12 tasks across the 30 concepts (Table 5). Each of the 30 concept tests consisted of 12 items. The girls attained higher mean scores than the boys on each of the 12 task attainment tests. For both the boys and the girls the first five tasks from the schema referred to previously (page 2) received the highest five ranks. They did not progress from one to five but the mean scores varied only slightly. For the boys it is of note that the rank orders of the mean scores for Tasks 6 and 7, selection of relevant and irrelevant attributes, were 10 and 12, respectively. Ranks of the other means appear to follow the general progression of the schema. For the girls it is of note that

Table 5
Means, Standard Deviations, and Hoyt Reliabilities for Tests of Task Attainment—Science

Task Number[a]	Mean		Standard Deviation		Hoyt Reliability	
	Boys	Girls	Boys	Girls	Boys	Girls
1	23.17	24.54	5.14	4.51	.84	.83
2	22.22	23.44	5.74	4.80	.87	.84
3	23.50	24.11	4.46	3.60	.80	.72
4	23.34	23.65	4.20	3.38	.76	.66
5	22.95	23.57	5.36	4.30	.85	.78
6	18.76	20.18	6.10	5.61	.85	.83
7	16.76	18.05	6.30	5.74	.85	.83
8	20.17	21.37	6.81	5.76	.89	.85
9	19.06	20.26	6.48	5.99	.87	.86
10	20.67	21.04	6.50	5.94	.88	.87
11	18.82	19.49	5.66	4.81	.83	.77
12	17.32	17.63	5.90	5.52	.83	.81

[a]1. Given name of attribute, select example of attribute.
2. Given example of attribute, select name of attribute.
3. Given name of concept, select example of concept.
4. Given name of concept, select nonexample of concept.
5. Given example of concept, select name of concept.
6. Given name of concept, select relevant attribute.
7. Given name of concept, select irrelevant attribute.
8. Given definition of concept, select name of concept.
9. Given name of concept, select definition of concept.
10. Given name of concept, select supraordinate concept.
11. Given name of concept, select subordinate concept.
12. Given names of two concepts, select principle relating them.

the rank order of the mean score on Task 7 was 11. All else seems to follow the general order of the schema.

Standard Deviations

The standard deviations for the tests of concept attainment were greater for the boys than for the girls on 26 of the 30 concept tests. In the biological science area, the standard deviations for boys were greater than those for girls on eight of the ten tests, and in both the earth science and the physical science areas, the standard deviations for boys were greater than those for girls on nine of the ten tests. The standard deviations on the task test scores were greater for the boys in all 12 instances.

Reliabilities

Hoyt reliabilities on the tests of concept attainment for the boys were greater than or equal to those for the girls on 23 of the 30 tests. In the earth science area, the test reliabilities for boys were greater than or equal to those for the girls on seven of the ten tests. In the biological and physical science areas, the test reliabilities for boys were greater than or equal to those for the girls on eight of the ten tests. On the task tests the reliabilities for the boys were greater than those for the girls in every instance, the largest difference in reliabilities being 0.10.

Item Indices

The item indices for the 360 items pre-

pared for use in this study (12 tasks, 30 concepts) are presented in Table 6. Again, the analyses for the boys and the girls are presented separately, and the information provided for each item includes the proportion correct, the biserial correlation, the X_{50}, and beta. Except for the proportion correct which is the same whether the item was part of the concept test score or the task test score, the item indices information is presented for the way in which the item performed as part of the concept test score or as part of the task test score. (The actual performances required by the respective tasks are listed at the bottom of Table 5.) Note that the item indices for the respective items are presented for the correct choice only. If one wishes to identify the specific concept and/or task for which data are presented, it is necessary to refer back to Tables 4 and 5. For example, Concept No. 2 as indicated in Table 6 would be Cell and the 1-12 tasks as indicated in Column 2 of Table 6 are for that respective concept. The third column in Table 6 in which the items are numbered consecutively from 1 to 360 is simply an ordering device and does not present any additional information regarding item indices on concept test scores or task test scores.

Item Difficulty

As indicated earlier in this paper, the mean scores on the tests indicated that the tests were more difficult for the boys than they were for the girls. This was evidenced by the fact that the girls attained higher mean scores on 26 of the 30 concept tests. Usually, however, it is not good practice to judge the difficulty of a test and its respective items on the basis of mean scores. More precise measures of item difficulty can be obtained from data regarding the proportion of items correct and the X_{50}.

As can be seen from Table 6, the girls attained a higher proportion of correct responses on a majority of the test items. On 22 of the 30 concept attainment tests, the girls had the highest proportion correct on more than 50% of the items, and on 16 of the 30 tests the proportion correct for the girls was higher than for boys on over two-thirds of the items.

The proportion correct as a means of judging item difficulty gives a more precise indication of item difficulty than a mean test score but is not as precise as the X_{50}. Further, the proportion correct does not give a good measure of the way in which an item is performing in reference to a criterion score. For instance, in this study, the proportion correct remains the same for both the criterion scores used, namely the concept attainment score and the task attainment score. The index tells how many students responded to the correct answer for an item but it says nothing about their ability level as measured by the specific criterion score.

A more precise measure of item difficulty than either the mean scores or proportion correct is the item difficulty index X_{50}. This difficulty index gives in standard deviation units the criterion score at which a subject would have a 50-50 chance of getting the item correct. For example, an X_{50} value of 1.50 for an item indicates that subjects with a criterion score 1.50 standard deviation units above the mean have a 50% chance of answering that item correctly. Subjects with a criterion score higher than this would have a greater chance of answering that item correctly and subjects with a criterion score lower than this would have a lesser chance. Likewise, an X_{50} value of -1.50 indicates that subjects with a criterion score 1.50 standard deviation units below the mean would have a 50% chance of getting that item correct. For a higher score the chance would be greater and for a lower score the chance would be less. Knowledge of both the X_{50} and the beta for a specific item permits one to easily determine the probability of answering an item correctly for any point on a criterion scale (Baker, 1964). It may be pointed out that when $p = .50$, $X_{50} = .00$; when p is greater than .50 then X_{50} will be negative; and for a certain p the higher the beta value the closer to zero will be the X_{50} value.

On the basis of the X_{50} scores for the items on the tests of concept attainment and task attainment, there is further support for the fact that the test items were more difficult for the boys than they were for the girls. The X_{50} values favored the girls on 248 of the 360 items when the concept score was the criterion score and on 239 of the items when the task score was the criterion score. Of note is that for 215 of these items the X_{50} item indices favor the girls on both the concept test score and the task test score.

Further information about the appropriateness of an item and its respective distractors can be obtained from an examination of the biserial correlation and beta. These indices are quite closely related since beta is computed as a function of the biserial correlation (Baker, 1969). However, the relationship is not linear. From .00 to about .30 (absolute) they are very nearly the same. Beyond this

(Continued on page 26)

Table 6
Item Indices Based on Concept and Task Criterion Scores—Science

			Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
Concept	Task	Item	Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
1	1	1	91	98	.64	.54	.20	.14	-2.14	-2.54	-11.00	-15.91	.83	.64	.20	.14
	2	2	89	94	.88	.90	.62	.75	-1.41	-1.38	-2.48	-2.06	1.84	2.03	.79	1.12
	3	3	91	90	.67	.55	.76	.53	-2.00	-2.41	-1.70	-2.45	.90	.67	1.19	.63
	4	4	82	83	.52	.48	.63	.43	-1.73	-1.88	-1.54	-2.27	.61	.55	.81	.47
	5	5	82	91	.78	.60	.42	.51	-1.16	-1.51	-3.12	-2.58	1.25	.75	.47	.60
	6	6	75	79	.48	.53	.51	.55	-1.42	-1.30	-1.60	-1.48	.55	.62	.59	.65
	7	7	61	70	.82	.76	.74	.59	-.35	-.38	-.70	-.88	1.42	1.17	1.12	.74
	8	8	70	79	.60	.37	.59	.50	-.90	-1.47	-1.37	-1.62	.74	.39	.74	.58
	9	9	46	47	.52	.42	.57	.42	.18	.23	.11	.15	.61	.46	.70	.47
	10	10	61	58	.48	.49	.58	.44	-.56	-.56	-.36	-.48	.55	.56	.71	.49
	11	11	84	89	.74	.68	.78	.72	-1.33	-1.46	-1.58	-1.72	1.11	.92	1.27	1.04
	12	12	55	63	.57	.42	.68	.51	-.21	-.29	-.49	-.65	.70	.46	.93	.59
2	1	13	48	48	.22	.03	.36	.20	.18	1.23	.15	.26	.23	.03	.39	.21
	2	14	61	69	.63	.71	.68	.70	-.45	-.40	-.72	-.69	.82	1.02	.93	.99
	3	15	82	87	.62	.67	.56	.46	-1.50	-1.39	-1.99	-2.42	.79	.89	.68	.52
	4	16	61	42	.60	.50	.48	.26	-.48	-.57	.40	.74	.75	.58	.54	.27
	5	17	65	72	.71	.55	.53	.57	-.52	-.68	-1.10	-1.03	1.02	.66	.63	.70
	6	18	77	76	.66	.52	.62	.54	-1.11	-1.41	-1.14	-1.32	.89	.61	.79	.64
	7	19	60	72	.55	.55	.69	.60	-.47	-.47	-.84	-.96	.65	.65	.96	.76
	8	20	46	39	.53	.36	.34	.21	.18	.26	.86	1.41	.62	.39	.36	.21
	9	21	54	53	.55	.31	.46	.29	-.17	-.31	-.18	-.28	.66	.33	.52	.30
	10	22	57	58	.60	.63	.63	.58	-.29	-.28	-.32	-.35	.74	.81	.82	.70
	11	23	56	47	.61	.45	.52	.42	-.27	-.36	.14	.17	.77	.50	.62	.46
	12	24	65	66	.50	.51	.55	.51	-.74	-.72	-.73	-.79	.58	.60	.67	.59
3	1	25	89	92	.81	.77	.47	.60	-1.53	-1.62	-3.04	-2.36	1.39	1.19	.53	.75
	2	26	87	94	.90	.83	.94	.81	-1.26	-1.36	-1.63	-1.91	2.06	1.50	2.61	1.36
	3	27	90	97	.80	.75	.83	1.02	-1.63	-1.73	-2.19	.00	1.32	1.14	1.49	.00
	4	28	94	97	.77	.42	.86	.73	-2.04	-3.70	-2.12	-2.50	1.19	.47	1.67	1.06
	5	29	90	97	.81	.68	.88	.82	-1.58	-1.91	-2.12	-2.28	1.46	.93	1.88	1.43
	6	30	65	82	.78	.71	.92	.81	-.50	-.54	-1.01	-1.14	1.24	1.02	2.28	1.39
	7	31	80	88	.72	.70	.82	.60	-1.14	-1.19	-1.45	-2.00	1.05	.97	1.46	.75
	8	32	72	75	.71	.61	.71	.45	-.83	-.95	-.96	-1.51	1.00	.78	1.01	.51
	9	33	73	78	.45	.43	.54	.57	-1.33	-1.38	-1.40	-1.33	.51	.48	.64	.69

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
			Boys	Girls	Boys		Girls		Boys		Girls		Boys		Girls	
					Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
3	10	34	52	48	.65	.61	.62	.51	-.06	-.07	.07	.09	.85	.76	.79	.59
	11	35	77	84	.58	-.23	.83	.55	-1.26	-7.79	-1.18	-1.79	.72	-.23	1.50	.66
	12	36	73	77	.81	.70	.80	.65	-.76	-.89	-.92	-1.13	1.36	.97	1.32	.85
4	1	37	73	72	-.82	.78	.69	.55	-.75	-.79	-.85	-1.08	1.43	1.24	.96	.65
	2	38	58	66	.65	.61	.65	.60	-.29	-.31	-.64	-.69	.86	.78	.86	.74
	3	39	84	89	.60	.59	.62	.59	-1.69	-1.71	-1.95	-2.05	.74	.73	.80	.74
	4	40	91	90	.49	.51	.61	.50	-2.72	-2.59	-2.12	-2.62	.56	.60	.78	.57
	5	41	72	81	.70	.64	.69	.49	-.83	-.91	-1.25	-1.78	.98	.83	.97	.56
	6	42	72	75	.63	.58	.81	.66	-.92	-1.02	-.81	-.99	.82	.70	1.40	.89
	7	43	72	79	.73	.63	.73	.60	-.80	-.93	-1.10	-1.34	1.06	.81	1.06	.75
	8	44	80	84	.89	.83	.83	.83	-.95	-1.02	-1.19	-1.19	1.92	1.50	1.48	1.48
	9	45	80	86	.67	.63	.81	.68	-1.15	-1.23	-1.35	-1.62	.90	.81	1.40	.93
	10	46	73	89	.83	.80	.81	.67	-.74	-.77	-1.53	-1.84	1.50	1.34	1.38	.91
	11	47	74	72	.71	.74	.83	.66	-.91	-.88	-.70	-.88	1.02	1.10	1.48	.87
	12	48	52	53	.56	.43	.53	.49	-.07	-.09	-.14	-.15	.68	.48	.63	.56
5	1	49	81	71	.49	.53	.69	.38	-1.77	-1.63	-.82	-1.48	.56	.63	.96	.41
	2	50	85	84	.55	.47	.52	.29	-1.89	-2.21	-1.93	-3.45	.65	.53	.61	.30
	3	51	43	40	.72	.42	.44	.25	.24	.42	.59	1.03	1.04	.47	.49	.26
	4	52	46	47	.68	.50	.62	.41	.16	.21	.13	.20	.92	.58	.79	.45
	5	53	72	70	.73	.66	.66	.42	-.80	-.88	-.81	-1.26	1.08	.89	.88	.47
	6	54	62	70	.85	.65	.73	.62	-.36	-.47	-.73	-.86	1.59	.85	1.07	.79
	7	55	65	75	.60	.53	.62	.60	-.64	-.73	-1.06	-1.10	.76	.62	.79	.75
	8	56	66	67	.73	.75	.66	.59	-.57	-.56	-.67	-.75	1.06	1.12	.88	.74
	9	57	59	59	.67	-.21	.72	.61	-.33	-3.44	-.33	-.40	.90	-.21	1.04	.76
	10	58	72	73	.73	.83	.62	.79	-.78	-.68	-1.00	-.79	1.06	1.51	.79	1.30
	11	59	60	59	.67	.50	.67	.39	-.38	-.52	-.33	-.57	.91	.57	.91	.42
	12	60	31	26	.38	.10	.53	.41	1.32	4.88	1.19	1.54	.42	.10	.63	.45
6	1	61	91	96	.46	.52	.79	.84	-2.89	-2.56	-2.22	-2.10	.52	.61	1.31	1.57
	2	62	84	87	.69	.90	.55	.56	-1.44	-1.10	-2.06	-2.03	.95	2.08	.66	.68
	3	63	89	89	.55	.56	.61	.65	-2.27	-2.21	-2.03	-1.91	.65	.68	.77	.85
	4	64	64	67	.71	.69	.62	.30	-.51	-.52	-.72	-1.51	1.00	.94	.78	.31
	5	65	67	62	.75	.65	.74	.42	-.57	-.66	-.40	-.71	1.15	.86	1.11	.46
	6	66	66	63	.55	.41	.55	.53	-.75	-1.01	-.62	-.65	.66	.45	.66	.62
	7	67	34	37	.39	.30	.55	.43	1.06	1.37	.60	.77	.43	.32	.66	.48

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
			Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
6	8	68	46	53	.74	.57	.72	.56	.13	.17	-.09	-.11	1.10	.70	1.05	.68
	9	69	49	49	.72	.49	.69	.43	.02	.03	.02	.03	1.04	.56	.95	.47
	10	70	85	93	.75	.79	.69	.82	-1.38	-1.31	-2.14	-1.80	1.13	1.30	.96	1.44
	11	71	69	75	.72	.76	.57	.56	-.68	-.64	-1.19	-1.22	1.05	1.17	.70	.68
	12	72	43	36	.54	.33	.36	.31	.33	.53	1.03	1.19	.64	.35	.38	.33
7	1	73	83	81	.75	.58	.84	.76	-1.29	1.66	-1.07	-1.18	1.14	.72	1.54	1.15
	2	74	80	82	.90	.89	.92	.85	-.94	-.95	-1.01	-1.09	2.10	1.98	2.32	1.58
	3	75	83	91	.75	.79	.69	.63	-1.29	-1.22	-1.91	-2.10	1.14	1.30	.96	.81
	4	76	83	85	.56	.69	.75	.72	-1.70	-1.37	-1.40	-1.46	.67	.95	1.14	1.04
	5	77	83	90	.59	.58	.72	.73	-1.64	-1.66	-1.75	-1.73	.73	.72	1.04	1.06
	6	78	58	67	.72	.57	.50	.49	-.28	-.36	-.90	-.91	1.04	.69	.57	.56
	7	79	70	66	.72	.65	.72	.66	-.73	-.81	-.56	-.61	1.03	.85	1.05	.87
	8	80	66	75	.85	.79	.89	.79	-.47	-.51	-.76	-.85	1.60	1.28	1.92	1.31
	9	81	67	73	.78	.72	.93	.86	-.57	-.62	-.65	-.70	1.25	1.03	2.46	1.71
	10	82	78	83	.84	.81	.88	.74	-.94	-.98	-1.08	-1.30	1.52	1.37	1.86	1.09
	11	83	74	77	.84	.77	.85	.73	-.77	-.85	-.87	-1.02	1.56	1.20	1.64	1.08
	12	84	68	75	.68	.57	.64	.61	-.70	-.84	-1.05	-1.10	.93	.69	.84	.77
8	1	85	90	93	.53	.72	.61	.61	-2.40	-1.76	-2.48	-2.47	.62	-1.05	.77	.77
	2	86	87	94	.79	.65	.88	.91	-1.43	-1.73	-1.79	-1.73	1.29	.87	1.84	2.15
	3	87	80	85	.80	.63	.98	.71	-1.04	-1.31	-1.07	-1.49	1.32	-.45	4.74	1.00
	4	88	81	93	.82	.55	.99	.55	-1.05	-1.58	-1.50	-2.67	1.44	.66	6.10	.67
	5	89	84	90	.86	.70	.95	.64	-1.18	-1.45	-1.34	-2.01	1.67	.98	3.10	.83
	6	90	69	81	.71	.61	.83	.64	-.71	-.83	-1.08	-1.40	1.00	.76	1.49	.83
	7	91	68	79	.81	.67	.82	.70	-.57	-.68	-.97	-1.15	1.37	.91	1.44	.97
	8	92	84	92	.86	.72	.94	.81	-1.18	-1.40	-1.48	-1.73	1.67	1.04	2.81	1.37
	9	93	82	84	.86	.59	.73	.48	-1.06	-1.54	-1.36	-2.04	1.66	.72	1.06	-.68
	10	94	84	84	.75	.54	.82	.76	-1.35	-1.88	-1.21	-1.30	1.13	.64	1.42	1.17
	11	95	70	82	.75	.58	.94	.60	-.69	-.90	-.97	-1.52	1.15	.71	2.64	.75
	12	96	82	86	.51	.47	.75	.67	-1.76	-1.94	-1.44	-1.62	.60	.53	1.14	.90
9	1	97	61	67	.57	.50	.68	.53	-.50	-.58	-.66	-.85	.69	.57	.92	.62
	2	98	70	70	.75	.68	.73	.64	-.70	-.77	-.73	-.83	1.13	.93	1.06	.84
	3	99	85	77	.61	.62	.72	.63	-1.69	-1.67	-1.04	-1.19	.77	.79	1.02	.80
	4	100	72	69	.31	.22	.62	.44	-1.82	-2.54	-.80	-1.14	.33	.23	.80	.49
	5	101	87	75	.52	.51	.57	.36	-2.16	-2.22	-1.16	-1.82	.62	.59	.69	.39

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
			Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
9	6	102	55	58	.54	.50	.57	.54	-.22	-.24	-.37	-.39	.65	.58	.70	.64
	7	103	44	49	.54	.51	.58	.55	.30	.32	.06	.06	.65	.59	.70	.65
	8	104	73	73	.74	.75	.65	.55	-.79	.76	-.96	-1.13	1.10	1.14	.85	.66
	9	105	61	60	.79	.75	.71	.66	-.36	-.38	-.37	-.40	1.30	1.14	1.00	.87
	10	106	81	80	.66	.67	.63	.67	-1.34	-1.33	-1.34	-1.25	.88	.89	.80	.90
	11	107	55	55	.74	.51	.60	.50	-.16	-1.70	-.20	-.24	1.09	-.59	.75	.58
	12	108	65	65	.57	.56	.56	.49	-.65	-.66	-.70	-.81	.70	.68	.69	.56
10	1	109	85	88	.64	.68	.80	-.66	-1.62	-1.52	-1.49	-2.56	.83	.92	1.34	-.87
	2	110	89	85	.42	.38	.56	.57	-2.98	-3.25	-1.82	-1.77	.46	.41	.67	.70
	3	111	65	71	.74	.48	.69	.48	-.51	-.78	-.78	-1.13	1.09	.55	.96	.56
	4	112	77	81	.67	.58	.69	.51	-1.10	-1.27	-1.30	-1.77	.90	.71	.95	.59
	5	113	63	64	.64	.60	.70	.63	-.52	-.55	-.50	-.56	.83	.74	.98	.80
	6	114	53	63	.57	.43	.61	.48	-.14	-.19	-.56	-.71	.70	.47	.77	.55
	7	115	57	57	.33	.41	.55	.44	-.53	-.43	-.31	-.39	.35	.44	.65	.48
	8	116	73	85	.76	.77	.96	.96	-.79	-.78	-1.08	-1.08	1.18	1.19	3.32	3.51
	9	117	81	80	.82	.74	.87	.83	-1.05	-1.17	-.98	-1.01	1.45	1.04	1.76	1.53
	10	118	65	76	.84	.67	.81	.70	-.44	-.56	-.89	-1.02	1.57	.90	1.37	.99
	11	119	60	67	.76	.71	.73	.76	-.32	-.35	-.60	-.57	1.16	1.00	1.06	1.16
	12	120	60	67	.71	.65	.88	.75	-.37	-.40	-.51	-.59	.99	.85	1.83	1.14
11	1	121	76	83	.65	.57	.52	.49	-1.11	-1.25	-1.85	-1.95	.85	.70	.60	.56
	2	122	35	24	.30	.19	.37	.21	1.26	1.93	1.92	3.40	.31	.20	.40	.21
	3	123	87	92	.68	.84	.64	.69	-1.67	-1.34	-2.24	-2.06	.92	1.56	.82	.96
	4	124	87	93	.79	.82	.78	.67	-1.44	-1.38	-1.94	-2.24	1.27	1.42	1.25	.91
	5	125	91	97	.53	.64	.74	.59	-2.52	-2.05	-2.44	-3.10	.62	.85	1.11	.72
	6	126	48	63	.72	.50	.57	.37	.08	.11	-.59	-.92	1.03	.57	.70	.40
	7	127	56	53	.70	.62	.62	.62	-.23	-.26	-.12	-.12	.99	.79	.78	.80
	8	128	65	80	.73	.72	.79	.79	-.53	-.54	-1.06	-1.07	1.06	1.04	1.30	1.27
	9	129	71	77	.77	.76	.78	.72	-.72	-.73	-.95	-1.04	1.19	1.17	1.26	1.04
	10	130	62	60	.79	.72	.72	.65	-.40	-.43	-.35	-.38	1.30	1.05	1.03	.85
	11	131	83	86	.80	.77	.44	.52	-1.21	-1.26	-2.50	-2.13	1.33	1.20	.49	.60
	12	132	59	58	.68	.72	.61	.49	-.32	-.30	-.35	-.43	.92	1.05	.76	.56
12	1	133	89	93	.73	.75	.77	.81	-1.70	-1.64	-1.93	-1.82	1.06	1.15	1.19	1.40
	2	134	80	84	.69	.64	.69	.72	-1.20	-1.30	-1.46	-1.39	.96	.83	.95	1.04
	3	135	84	91	.89	.79	.70	.61	-1.13	-1.28	-1.89	-2.19	2.00	1.28	.98	.76

Table 6 (Continued)

			Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
Concept	Task	Item	Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
12	4	136	84	87	.85	.69	.65	.59	-1.17	-1.43	-1.72	-1.91	1.59	.96	.86	.72
	5	137	84	91	.78	.61	.70	.65	-1.30	-1.67	-1.89	-2.04	1.23	.76	.98	.85
	6	138	54	51	.66	.57	.64	.47	-.16	-.19	-.04	-.05	.88	.69	.83	.53
	7	139	48	44	.58	.40	.59	.42	.09	.13	.24	.34	.72	.44	.73	.46
	8	140	84	81	.81	.80	.71	.68	-1.22	1.23	-1.17	-1.28	1.40	1.35	1.10	.92
	9	141	68	76	.73	.57	.79	.61	-.63	-.81	-.91	-1.18	1.07	.69	1.28	.77
	10	142	58	49	.63	.48	.59	.45	-.30	-.39	.02	.03	.82	.55	.76	.50
	11	143	62	76	.72	.72	.63	.61	-.44	-.44	-1.14	-1.18	1.03	1.03	.81	.77
	12	144	73	75	.68	.55	.64	.59	-.88	-1.09	-1.06	-1.15	.94	.66	.82	.72
13	1	145	85	86	.68	.71	.63	.60	-1.56	-1.48	-1.76	-1.84	.92	1.02	.80	.75
	2	146	83	90	.62	.56	.65	.82	-1.51	-1.70	-1.93	-1.54	.80	.67	.86	1.42
	3	147	48	50	.29	.20	.25	.01	.19	.28	.02	.59	.30	.20	.25	.01
	4	148	90	92	.58	.56	.51	.42	-2.24	-2.31	-2.73	-3.36	.71	.68	.59	.46
	5	149	83	90	.88	.88	.77	.58	-1.07	-1.08	-1.66	-2.22	1.87	1.85	1.20	.70
	6	150	74	75	.65	.55	.72	.55	-.98	-1.15	-.91	-1.20	.85	.66	1.05	.65
	7	151	54	54	.56	.44	.54	.51	-.17	-.22	-.19	-.20	.68	.48	.65	.59
	8	152	70	81	.80	.77	.77	.78	-.67	-.70	-1.15	-1.14	1.33	1.20	1.20	1.23
	9	153	72	83	.79	.87	.96	.95	-.72	-.65	-.98	-.99	1.29	1.79	3.42	3.00
	10	154	80	80	.82	-.49	.88	.74	-1.00	-3.76	-.97	-1.16	1.45	-.57	1.81	1.09
	11	155	78	83	.80	.69	.75	.63	-.99	-1.14	-1.29	-1.54	1.31	.97	1.14	.81
	12	156	65	72	.68	.63	.53	.49	-.55	-.59	-1.09	-1.17	.92	.82	.63	.57
14	1	157	57	73	.65	.53	.68	.65	-.27	-.33	-.90	-.94	.86	.62	.93	.86
	2	158	61	61	.66	-.32	.61	.45	-.41	-3.84	-.44	-.60	.88	-.33	.77	.50
	3	159	88	87	.52	.69	.49	.56	-2.21	-1.68	-2.27	-2.01	.61	.95	.57	.67
	4	160	88	92	.54	.50	.52	.47	-2.13	-2.33	-2.75	-3.00	.65	.57	.61	.54
	5	161	72	70	.61	.58	.69	.51	-.95	-1.01	-.74	-1.00	.78	.71	.95	.59
	6	162	60	67	.63	.48	.76	.65	-.41	-.54	-.58	-.68	.80	.54	1.18	.86
	7	163	66	66	.53	.39	.57	.46	-.78	-1.08	-.73	-.90	.63	.42	.69	.51
	8	164	80	86	.70	.74	.85	.95	-1.20	-1.14	-1.28	-1.15	.99	1.10	1.59	2.89
	9	165	78	86	.81	.83	.85	.85	-.98	-.95	-1.30	-1.30	1.36	1.51	1.61	1.59
	10	166	58	60	.55	.48	.72	.60	-.37	-.42	-.36	-.44	.65	.55	1.05	.74
	11	167	68	66	.54	.50	.42	.38	-.86	-.93	-.95	-1.05	.64	.57	.47	.41
	12	168	56	55	.66	.56	.58	.56	-.23	-.27	-.23	-.23	.88	.68	.70	.67

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
			Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
15	1	169	62	70	.73	.70	.55	.58	-.43	-.45	-.97	-.91	1.06	.98	.66	.72
	2	170	73	85	.82	.85	.64	.73	-.76	-.72	-1.64	-1.44	1.40	1.62	.83	1.07
	3	171	62	56	.70	.47	.58	.45	-.45	-.68	-.26	-.33	.99	.53	.71	.50
	4	172	90	82	.38	.32	.33	.41	-3.30	-3.96	-2.81	-2.26	.42	.34	.35	.45
	5	173	76	73	.73	.63	.68	.48	-.96	-1.12	-.92	-1.29	1.07	.80	.92	.55
	6	174	64	30	.48	.29	.44	.31	.92	1.54	1.16	1.63	.55	.30	.49	.33
	7	175	60	62	.67	.63	.61	.58	-.38	.41	-.49	-.52	.91	.80	.77	.71
	8	176	68	78	.73	.86	.81	.85	-.67	-.57	-.94	-.89	1.08	1.66	1.39	1.65
	9	177	68	75	.72	.71	.78	.68	-.66	-.67	-.87	-.99	1.04	1.02	1.23	.92
	10	178	55	55	.48	.45	.49	.40	-.28	-.30	-.27	-.32	.55	.50	.56	.44
	11	179	52	37	.44	.33	.45	.30	-.12	-.16	.71	1.08	.49	.35	.51	.31
	12	180	63	63	.86	.75	.83	.82	-.40	-.46	-.39	-.39	1.67	1.13	1.46	1.43
16	1	181	76	87	.73	.80	.82	.75	-.98	-.90	-1.40	-1.53	1.08	1.33	1.41	1.12
	2	182	77	83	.78	-.43	.71	.70	-.97	-2.80	-1.34	-1.36	1.24	-.48	1.02	.98
	3	183	90	86	.67	.65	.58	.67	-1.93	-1.99	-1.86	-1.61	.91	.86	.72	.91
	4	184	90	87	.49	.43	.57	.49	-2.64	-3.04	-1.96	-2.26	.57	.47	.70	.57
	5	185	74	64	.82	.70	.71	-.54	-.77	-.91	-.52	-3.20	1.45	.97	1.02	-.64
	6	186	82	74	.84	.63	.79	.64	-1.10	-1.47	-.80	-.99	1.54	.81	1.28	.84
	7	187	58	50	.73	.72	.81	.65	-.28	-.28	-.01	-.01	1.06	1.05	1.40	.86
	8	188	68	64	.80	.76	.76	.60	-.58	-.61	-.47	-.60	1.33	1.17	1.18	.75
	9	189	61	68	.77	.64	.82	.75	-.37	-.45	-.57	-.62	1.22	.82	1.46	1.15
	10	190	63	69	.74	.70	.79	.77	-.47	-.49	-.62	-.63	1.08	.97	1.30	1.20
	11	191	48	41	.52	.36	.59	.56	.10	.15	.41	.43	.60	.39	.73	.67
	12	192	68	61	.69	.70	.78	.75	-.69	-.68	-.37	-.39	.95	.99	1.25	1.12
17	1	193	84	88	.61	.65	.81	.78	-1.65	-1.56	-1.47	-1.53	.77	.85	1.40	1.26
	2	194	69	76	.77	.68	.75	.70	-.64	-.72	-.93	-.99	1.19	.92	1.13	.98
	3	195	83	87	.77	.68	.60	.60	-1.22	-1.40	-1.88	-1.86	1.22	.92	.74	.76
	4	196	90	91	.59	.72	.50	.31	-2.20	-1.81	-2.68	-4.41	.73	1.04	.58	.32
	5	197	86	87	.60	-.69	.84	.86	-1.81	-2.25	-1.33	-1.31	.75	-.96	1.56	1.67
	6	198	72	72	.67	.64	.69	.64	-.85	-.89	-.86	-.92	.90	.83	.95	.84
	7	199	60	68	.66	.55	.82	.69	-.39	-.47	-.58	-.69	.87	.66	1.44	.95
	8	200	67	63	.60	.58	.66	.58	-.72	-.74	-.50	-.57	.75	.72	.89	.71
	9	201	72	66	.75	.68	.65	.50	-.76	-.84	-.65	-.85	1.13	.93	.85	.57
	10	202	65	70	.74	-.41	.73	.74	-.50	-2.90	-.72	-.72	1.11	-.45	1.08	1.10

Table 6 (Continued)

			Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
Concept	Task	Item	Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
17	11	203	22	31	.30	-.37	.14	.05	2.59	-1.85	3.54	9.44	.32	-.40	.15	.05
	12	204	63	68	.67	.71	.72	.65	-.51	-.49	-.63	-.70	.91	1.00	1.04	.87
18	1	205	94	98	-.64	.43	.32	.61	-2.50	-3.52	-6.15	-3.25	-.84	.48	.34	.78
	2	206	85	93	.62	.59	.76	.87	-1.66	-1.77	-1.99	-1.73	.79	.72	1.17	1.78
	3	207	69	78	.79	.59	.87	.63	-.64	-.85	-.89	-1.23	1.27	.73	1.76	.81
	4	208	74	78	.68	.42	.82	.54	-.93	-1.49	-.96	-1.44	.92	.47	1.42	.65
	5	209	58	64	.68	.45	.72	.51	-.28	-.42	-.52	-.72	.93	.50	1.03	.60
	6	210	63	70	.68	.65	.71	.58	-.48	-.51	-.75	-.92	.93	.86	1.00	.71
	7	211	54	69	.65	.66	.63	.51	-.15	-.14	-.77	-.95	.86	.88	.81	.60
	8	212	65	61	.57	.38	.76	.58	-.67	-1.01	-.38	-.50	.70	.41	1.18	.72
	9	213	56	71	.74	.72	.77	.58	-.20	-.21	-.73	-.98	1.11	1.04	1.21	.71
	10	214	74	82	.79	.71	.67	.65	-.80	-.90	-1.37	-1.41	1.31	1.00	.91	.86
	11	215	42	41	.42	.37	.63	.36	.48	.55	.38	.66	.46	.40	.81	.39
	12	216	59	68	.63	.55	.47	.48	-.35	-.39	-1.01	-.99	.81	.66	.53	.55
19	1	217	86	88	.64	.62	.67	.63	-1.70	-1.74	-1.77	-1.86	.82	.80	.89	.82
	2	218	72	77	.66	.65	.64	.43	-.88	-.90	-1.15	-1.70	.89	.85	.83	.48
	3	219	90	94	.94	.96	.98	.85	-1.35	-1.33	-1.60	-1.85	2.82	3.28	5.08	1.61
	4	220	91	94	.43	.55	.48	.09	-3.13	-2.44	-3.21	-17.40	.47	.65	.55	.09
	5	221	67	67	.55	.47	.45	.41	-.78	-.92	-.96	-1.05	.66	.53	.51	.45
	6	222	70	74	.73	.77	.79	.70	-.72	-.67	-.82	-.92	1.06	1.22	1.29	.99
	7	223	66	76	.67	.55	.61	.52	-.62	-.75	-1.13	-1.33	.90	.67	.78	.61
	8	224	75	83	.77	.79	.81	.69	-.88	-.86	-1.19	-1.41	1.22	1.30	1.39	.95
	9	225	63	64	.58	.51	.52	.48	-.59	-.68	-.69	-5.59	.72	.59	.61	-.16
	10	226	87	91	.62	.65	.50	.58	-1.82	-1.73	-2.65	-2.29	.79	.86	.58	.71
	11	227	58	64	.56	.57	.73	.63	-.34	-.34	-.51	-.59	.68	.69	1.05	.81
	12	228	54	55	.51	.47	.52	.47	-.18	-.20	-.25	-.28	.60	.53	.61	.54
20	1	229	72	81	.69	.73	.69	.57	-.85	-.80	-1.25	-1.51	.95	1.07	.96	.70
	2	230	81	87	.70	.74	.52	.50	-1.24	-1.18	-2.18	-2.28	.98	1.09	.61	.58
	3	231	80	91	.57	.63	.55	.57	-1.49	-1.34	-2.46	-2.36	.69	.82	.66	.70
	4	232	71	73	.39	.25	.29	.22	-1.43	-2.25	-2.10	-2.72	.42	.25	.30	.23
	5	233	88	92	.69	.74	.66	.57	-1.67	-1.57	-2.13	-2.44	.97	1.09	.87	.70
	6	234	81	86	.77	.65	.86	.70	-1.13	-1.33	-1.26	-1.56	1.20	.86	1.72	.97
	7	235	51	63	.55	.51	.71	.57	-.05	-.05	-.48	-.60	.67	.59	1.01	.69
	8	236	60	64	.71	.67	.76	.61	-.34	-.37	-.47	-.59	1.01	.90	1.18	.77

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
			Boys	Girls	Boys		Girls		Boys		Girls		Boys		Girls	
					Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
20	9	237	71	80	.74	.70	.85	.87	-.75	-.79	-.97	-.95	1.09	-.98	1.63	1.78
	10	238	77	79	.79	.76	.85	.68	-.95	-.99	-.95	-1.19	1.29	1.17	1.63	.94
	11	239	77	88	.72	.66	.75	.75	-1.03	-1.11	-1.57	-1.56	1.02	.89	1.13	1.14
	12	240	68	73	.75	.69	.65	.60	-.61	-.67	-.93	-1.01	1.15	.94	.85	.74
21	1	241	81	85	.63	.78	.60	.77	-1.38	-1.11	-1.68	-1.32	.80	1.25	.76	1.22
	2	242	72	80	.70	.74	.51	.56	-.81	-.77	-1.67	-1.52	.98	1.09	.59	.68
	3	243	46	37	.52	.33	.63	.32	.18	.29	.53	1.04	.60	.35	.80	.33
	4	244	43	31	.57	.36	.59	.39	.31	.48	.85	1.30	.70	.39	.73	.42
	5	245	49	47	.58	.44	.59	.40	.02	.03	.11	.16	.72	.49	.73	.43
	6	246	63	59	.57	.51	.60	-.11	-.61	-.67	-.37	-5.65	.69	.60	.74	-.12
	7	247	46	44	.70	.59	.50	.51	.14	.16	.28	.28	.96	.74	.59	.59
	8	248	66	60	.63	.64	.63	.54	-.66	-.65	-.39	-.46	.81	.82	.81	.64
	9	249	51	41	.65	.54	.85	.62	-.04	-.05	.27	.37	.86	.64	1.59	.80
	10	250	48	37	.51	.32	.54	.34	.11	.17	.61	.98	.59	.34	.65	.36
	11	251	52	49	.78	.58	.58	.38	-.07	-.09	.04	.06	1.23	.71	.70	.41
	12	252	44	34	.38	.36	.50	.36	.39	.42	.82	1.15	.41	.38	.58	.39
22	1	253	80	79	.62	.63	.58	.56	-1.37	-1.34	-1.40	-1.45	.78	.81	.72	.68
	2	254	76	81	.67	.61	.51	.48	-1.08	-1.18	-1.72	-1.83	.90	.77	.60	.55
	3	255	67	70	.59	.49	.55	.46	-.73	-.88	-.93	-1.10	.73	.56	.65	.52
	4	256	82	87	.67	.61	.71	.70	-1.35	-1.48	-1.58	-1.59	.91	.78	1.00	.99
	5	257	65	58	.56	.46	.68	.52	-.66	-.81	-.28	-.37	.68	.51	.93	.60
	6	258	49	56	.65	.63	.60	.53	.04	.04	-.27	-.30	.85	.81	.75	.62
	7	259	56	59	.69	.67	.60	.60	-.22	-.22	-.38	-.38	.96	.91	.75	.77
	8	260	64	68	.64	.64	.63	.61	-.56	-.56	-.74	-.76	.83	.84	.82	.78
	9	261	70	69	.55	.47	.67	.62	-.95	-1.11	-.75	-.81	.66	.53	.89	1.52
	10	262	74	75	.72	.72	.74	.84	-.90	-.90	-.91	-.80	1.05	1.04	1.09	.87
	11	263	56	68	.69	.74	.62	.66	-.22	-.20	-.74	-.69	.96	1.10	.78	.87
	12	264	61	59	.63	.59	.57	.74	-.43	-.46	-.42	-.51	.82	.73	.70	.54
23	1	265	51	60	.68	.57	.82	.73	-.02	-.02	-.31	-.36	.92	.69	1.46	1.07
	2	266	68	69	.73	.63	.76	.69	-.65	-.75	-.64	-.70	1.06	.81	1.17	.95
	3	267	81	75	.54	.45	.58	.44	-1.61	-1.93	-1.17	-1.52	.64	.50	.70	.49
	4	268	69	80	.58	.58	.70	.69	-.87	-.88	-1.21	-1.24	.71	.70	.99	.95
	5	269	83	86	.65	.70	.76	.80	-1.49	-1.38	-1.44	-1.35	.85	.98	1.15	1.34

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
			Boys	Girls	Boys		Girls		Boys		Girls		Boys		Girls	
					Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
23	6	270	56	58	.79	.73	.74	.66	-.20	-.22	-.27	-.30	1.30	1.06	1.10	.88
	7	271	52	49	.56	.60	.66	.58	-.07	-.07	.02	.03	.68	.76	.88	.71
	8	272	55	65	.80	.73	.71	.64	-.17	-.18	-.55	-.61	1.31	1.08	1.01	.83
	9	273	66	69	.81	.70	.81	.72	-.51	-.59	-.62	-.70	1.38	.99	1.36	1.03
	10	274	69	73	.74	.77	.66	.62	-.69	-.66	-.93	-.99	1.09	1.20	.87	.79
	11	275	56	58	.48	.35	.43	.25	-.31	-.43	-.45	-.76	.54	.37	.47	.26
	12	276	44	38	.23	.25	.36	.32	.64	.60	-1.65	.97	.24	.26	-.34	.34
24	1	277	90	91	.57	.56	.70	.90	-2.30	-2.31	-1.91	-1.50	.69	.68	.99	2.03
	2	278	81	83	.67	.66	.51	.71	-1.32	-1.34	-1.89	-1.35	.90	.88	.59	1.00
	3	279	81	78	.75	.36	.75	.68	-1.17	-2.46	-1.01	-1.12	1.15	.39	1.14	.93
	4	280	52	44	.56	.52	.51	.39	-.07	-.08	-1.20	.39	.68	.61	-.28	.42
	5	281	77	79	.63	.65	.68	.57	-1.16	-1.13	-1.20	-1.43	.82	.86	.92	.69
	6	282	42	46	.49	.47	.29	.16	.41	.43	.38	.71	.57	.54	.30	.16
	7	283	48	48	.43	.28	.29	.15	.13	.20	.22	.41	.48	.29	.30	.16
	8	284	74	70	.72	.69	.65	.46	-.90	-.94	-.80	-1.13	1.04	.96	.86	.52
	9	285	46	42	.61	.50	.50	.39	.18	.22	.40	.51	.77	.57	.57	.42
	10	286	66	62	.69	.73	.66	.57	-.58	-.55	-.45	-.52	.95	1.07	.88	.70
	11	287	76	70	.74	.60	.70	.69	-.98	-1.20	-.76	-.77	1.09	.75	.97	.95
	12	288	36	37	.33	.23	.38	.25	1.09	1.54	.87	1.31	.35	.24	.41	.26
25	1	289	85	88	.91	.73	.97	.84	-1.16	-1.45	-1.19	-1.38	2.25	1.06	4.33	1.52
	2	290	72	77	.76	.55	.74	.67	-.76	-1.06	-.99	-1.09	1.19	.66	1.11	.91
	3	291	94	95	.52	.73	.63	.82	-3.00	-2.14	-2.59	-1.99	.61	1.07	.82	1.45
	4	292	87	92	.54	.55	.70	.69	-2.11	-2.06	-1.99	-2.03	.64	.66	.99	.95
	5	293	80	79	.57	.55	.45	.35	-1.50	-1.54	-1.79	-2.32	.69	.65	.51	.37
	6	294	69	77	.70	.60	.81	.66	-.73	-.85	-.90	-1.10	.97	.74	1.40	.89
	7	295	56	63	.73	.61	.69	.53	-.20	-.24	-.50	-.64	1.07	.77	.94	.63
	8	296	67	69	.61	.50	.61	.53	-.70	-.86	-.80	-.92	.77	.58	.77	.63
	9	297	46	56	.68	.50	.62	.45	.16	.21	-.23	-.31	.92	.58	.78	.51
	10	298	78	70	.52	-.39	.70	.66	-1.51	-3.67	-.76	-.81	.61	-.43	.97	.87
	11	299	81	87	.60	.48	.72	.70	-1.44	-1.82	-1.56	-1.60	.75	.54	1.03	.98
	12	300	74	70	.61	.61	.57	.59	-1.04	-1.04	-.93	-.90	.77	.76	.70	.74
26	1	301	70	77	.67	.53	.67	.60	-.77	-.98	-1.09	-1.23	.91	.63	.91	.74
	2	302	42	37	.29	.21	.48	.41	.70	-8.42	-.69	.80	.30	-.22	.55	.45
	3	303	92	96	.59	.82	.81	.65	-2.43	-1.76	-2.13	-2.63	.73	1.41	1.38	.87

Table 6 (Continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
			Boys		Boys		Girls		Boys		Girls		Boys		Girls	
			Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
26	4	304	89	96	.63	.78	.73	.86	-1.97	-1.60	-2.36	-2.00	.81	1.23	1.07	1.69
	5	305	90	94	.79	.81	.74	.74	-1.60	-1.56	-2.07	-2.09	1.30	1.39	1.12	1.09
	6	306	33	46	.45	.41	.57	.53	.95	1.05	.16	.17	.51	.45	.70	.62
	7	307	56	66	.68	.70	.67	.72	-.22	-.21	-.61	-.57	.93	.97	.91	1.05
	8	308	65	65	.74	.60	.68	.57	-.50	-.62	-.58	-.69	1.11	.75	.93	.69
	9	309	67	81	.72	.60	.71	.70	-.60	-.71	-1.26	-1.27	1.05	.76	1.01	.99
	10	310	74	79	.75	.71	.71	.65	-.7	-2.86	-1.14	-1.25	1.12	-.51	1.02	.85
	11	311	59	64	.68	.57	.57	.54	-.34	-.41	-.65	.69	.92	.69	.69	.64
	12	312	38	35	.58	.55	.55	.48	.54	.58	.71	.82	.72	.65	.67	.55
27	1	313	69	80	.59	.63	.61	.68	-.83	-.78	-1.35	-1.22	.73	.81	.77	.92
	2	314	77	82	.66	.59	.56	.61	-1.14	-1.27	-1.62	-1.49	.87	.73	.68	.77
	3	315	56	56	.71	.37	.65	.43	-.21	-.40	-.22	-.32	1.01	.40	.86	.48
	4	316	53	54	.52	.33	.61	.40	-.13	-.20	-.15	-.23	.61	.34	.77	.44
	5	317	56	57	.63	.54	.62	.51	-.26	-.30	-.29	-.35	.80	.65	.79	.60
	6	318	69	82	.74	.67	.76	.75	-.69	-.75	-1.19	-1.21	1.09	.91	1.18	1.14
	7	319	40	42	.57	.39	.43	.32	.43	.62	.49	.66	.70	.43	.48	.33
	8	320	60	66	.42	.45	.49	.39	-.61	-.58	-.87	-1.10	.47	.50	.56	.42
	9	321	46	44	.41	.30	.30	.22	.23	.31	.53	.74	.44	.32	.31	.22
	10	322	56	51	.56	.59	.58	.53	-.29	-.27	-.06	-.06	.67	.74	.72	.62
	11	323	25	28	.32	.21	.37	.20	2.06	3.22	1.57	2.84	.34	.21	.39	.21
	12	324	53	58	.50	.48	.60	.42	-.16	-.17	-.33	-.48	.57	.55	.75	.46
28	1	325	63	66	.65	.58	.58	.54	-.50	-.57	-.70	-.74	.86	.71	.72	.65
	2	326	67	76	.76	.59	.78	.73	-.57	-.73	-.89	-.95	1.16	.73	1.26	1.07
	3	327	88	93	.69	.74	.67	.54	-1.70	-1.59	-2.16	-2.69	.97	1.11	.91	.64
	4	328	88	90	.85	.86	.68	.53	-1.39	-1.38	-1.86	-2.35	1.63	1.67	.92	.63
	5	329	82	92	.71	.65	.87	.90	-1.30	-1.42	-1.63	-1.59	1.01	.86	1.78	2.02
	6	330	75	80	.75	.73	.74	.60	-.91	-.93	-1.12	-1.40	1.15	1.08	1.13	.75
	7	331	62	79	.69	.64	.68	.49	-.43	-.47	-1.19	-1.67	.96	.83	.94	.56
	8	332	63	80	.80	.70	.58	.48	-.43	-.49	-1.45	-1.75	1.35	.99	.71	.55
	9	333	45	56	.72	.62	.72	.62	.19	.22	-.19	-.23	1.02	.80	1.05	.80
	10	334	65	70	.65	.67	.80	.66	-.59	-.58	-.65	-.80	.86	.91	1.31	.87
	11	335	83	92	.71	.52	.74	.54	-1.34	-1.81	-1.93	-2.65	1.00	.61	1.10	.64
	12	336	76	83	.73	.56	.65	.50	-.99	-1.29	-1.50	-1.92	1.06	.67	.85	.58

Table 6 (Continued)

			Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys		Girls		Boys		Girls		Boys		Girls	
Concept	Task	Item	Boys	Girls	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task	Concept	Task
29	1	337	75	85	.75	-.41	.80	.80	-.91	-3.18	-1.27	-1.27	1.14	-.45	1.33	1.33
	2	338	81	90	.79	.74	.90	.86	-1.12	-1.19	-1.44	-1.52	1.28	1.12	2.08	1.66
	3	339	82	90	.61	.59	.70	.71	-1.52	-1.57	-1.85	-1.84	.76	.73	.99	1.00
	4	340	88	87	.67	.40	.55	.61	-1.72	-2.89	-2.06	-1.86	.91	.44	.67	.77
	5	341	87	90	.72	.75	.82	.87	-1.58	-1.51	-1.54	-1.44	1.02	1.13	1.41	1.77
	6	342	70	73	.69	.65	.67	.64	-.76	-.81	-.90	-.94	.94	.85	.89	.84
	7	343	32	34	.50	.36	.65	.48	.92	1.27	.65	.89	.58	.39	.86	.54
	8	344	58	68	.77	.61	.72	.56	-.26	-.33	-.65	-.84	1.22	.78	1.03	.67
	9	345	62	70	.73	.66	.81	.76	-.41	-.45	-.65	-.70	1.07	.88	1.40	1.15
	10	346	63	64	.56	-.28	.58	.46	-.59	-4.97	-.61	-.77	.67	-.29	.70	.51
	11	347	77	71	.51	.49	.53	.38	-1.49	-1.54	-1.07	-1.48	.59	.56	.63	.41
	12	348	40	32	.57	.59	.38	.42	.46	.44	1.21	1.09	.69	.72	.41	.46
30	1	349	77	88	.85	.80	.80	.91	-.88	-.94	-1.45	-1.28	1.64	1.34	1.33	2.16
	2	350	79	85	.96	.82	.75	.74	-.84	-.98	-1.37	-1.38	3.64	1.45	1.12	1.10
	3	351	88	95	.60	.78	.66	.68	-1.96	-1.52	-2.48	-2.42	.76	1.24	.89	.92
	4	352	80	80	.78	.78	.61	.63	-1.06	-1.06	-1.39	-1.33	1.26	1.24	.76	.81
	5	353	83	90	.87	.83	.71	.94	-1.09	-1.14	-1.81	-1.36	1.77	1.48	.99	2.70
	6	354	59	63	.58	.37	.50	.41	-.40	-.62	-.66	.80	.71	.40	.57	.45
	7	355	44	46	.49	.53	.49	.34	.31	.28	.23	.32	.56	.63	.56	.37
	8	356	66	63	.55	.43	.58	.30	-.73	-.94	-.59	-1.14	.66	.47	.71	.31
	9	357	78	80	.85	-.58	.70	.44	-.93	-2.68	-1.21	-1.94	1.62	.72	.99	.49
	10	358	86	83	.74	.72	.70	.77	-1.46	-1.50	-1.38	-1.27	1.10	1.04	.99	1.19
	11	359	47	41	.48	.43	.48	.34	.14	.16	.48	.68	.54	.48	.54	.36
	12	360	47	54	.53	.62	.52	.45	.15	.13	-.20	-.23	.63	.78	.61	.50

point beta begins to increase quite rapidly in magnitude in comparison to the biserial correlation. It may be pointed out that beta is always equal to or greater (absolute) than the biserial correlation. As a general rule, .30 is often used as a lower cutting point for a desirable biserial correlation or a beta. For a total score composed of relatively few items, as is the concept score, a much higher minimum would be desirable.

It can be noted from Table 6 that only 36 of the 360 items (10%) had betas less than 0.30 when functioning as either part of a concept criterion score or a task criterion score. It is of further note that when raising the beta level to 0.40 only one-sixth of the test items have beta scores less than this value when functioning as either part of a concept criterion score or a task criterion score. For those items which had betas less than 0.40, 52 were weaker items when they functioned as part of a task criterion score. A comparison of items for the girls and the boys indicates no preference for one group over the other in terms of the magnitude of the betas.

In general, only 10 of the 360 items functioned poorly as part of both the concept criterion score and the task criterion score.

GPO 827-451-3

IV
Conclusions

The discussion of the means, standard deviations, and Hoyt reliability estimates presented in this paper for the concept attainment tests and the task attainment tests indicates that the tests as constructed are appropriate for facilitating the overall objectives of the project entitled "A Structure of Concept Attainment Abilities." Further support for the appropriateness of these test instruments is obtained from the examination of four different item indices: proportion correct, item criterion biserial correlation, X_{50}, and beta. The analyses of these four indices based on each of two criterion scores indicate that it is appropriate to use these tests to continue pursuance of the study of the relationship between concept attainment, task performance, and selected measures of cognitive ability. The item indices show that these items are of appropriate difficulty level for use with subjects at this age level. It is further indicated by analyses of these data that the majority of the items have desirable levels of discrimination indices when the item is both a part of a concept criterion score and a task criterion score.

References

Baker, F. B. An intersection of test score interpretation and item analysis. Journal of Educational Measurement, 1964, 1, 23-28.

Baker, F. B. FORTAP: A FORTRAN test analysis package. Department of Educational Psychology, The University of Wisconsin, 1969.

Bourne, L. E., Jr. Human conceptual behavior. Boston: Allyn & Bacon, 1966.

Frayer, D. A., Fredrick, W. C., & Klausmeier, H. J. A schema for testing the level of concept mastery. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 16, 1969.

Harris, M. L., Harris, C. W., Frayer, D. A., & Quilling, M. R. A structure of concept attainment abilities: The problem and strategies for attacking it. Wisconsin Research and Development Center for Cognitive Learning, Theoretical Paper No. 32, in press.

Klausmeier, H. J., Harris, C. W., Davis, J. K., Schwenn, E., & Frayer, D. A. Strategies and cognitive processes in concept learning. The University of Wisconsin, Cooperative Research Project No. 2850, 1968.

Voelker, A. M., Sorenson, J. S., & Frayer, D. A. An analysis of selected classificatory science concepts in preparation for writing tests of concept attainment. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 57, 1971.

Voelker, A. M., & Sorenson, J. S. Items for measuring the level of attainment of selected classificatory science concepts by intermediate grade children. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 58, 1971.

GPO 827-451-2

National Evaluation Committee

Helen Bain
Immediate Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
Universiy of Colorado

Jeanne S. Chall
Graduate School of Education
Harvard University

Francis S. Chase
Department of Education
University of Chicago

George E. Dickson
College of Education
University of Toledo

Hugh J. Scott
Superintendent of Public Schools
District of Columbia

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Benton J. Underwood
Department of Psychology
Northwestern University

Robert J. Wisner
Mathematics Department
New Mexico State University

Executive Committee

William R. Bush
Director of Program Planning and Management
and Deputy Director, R & D Center

Herbert J. Klausmeier, Committee Chairman
Director, R & D Center

Wayne Otto
Principal Investigator
R & D Center

Robert G. Petzold
Professor of Music
University of Wisconsin

Richard A. Rossmiller
Professor of Educational Administration
University of Wisconsin

James E. Walter
Coordinator of Program Planning
R & D Center

Russell S. Way, ex officio
Program Administrator, Title III ESEA
Wisconsin Department of Public Instruction

Faculty of Principal Investigators

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
Center Director
V. A. C. Henmon Professor
Educational Psychology

Stephen J. Knezevich
Professor
Educational Administration

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

Wayne Otto
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Peter A. Schreiber
Assistant Professor
English

Richard L. Venezky
Associate Professor
Computer Science

Alan M. Voelker
Assistant Professor
Curriculum and Instruction

Larry M. Wilder
Assistant Professor
Communication Arts